

IN THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application. An identifier indicating the status of each claim is provided.

Listing of Claims:

1. (Currently Amended) A speech recognition apparatus in which a score reflecting an acoustic likelihood of the results of speech recognition of an input speech is calculated and in which the speech is recognized based on the score, comprising:

input means to receive the input speech;

extraction means for extracting characteristic values of said input speech, the input speech comprising a plurality of input words;

characteristic value storage means for storing characteristic values and an extraction time point associated with each characteristic value;

word concatenation information storage means for storing word concatenation information that is the relation of input words of a word sequence representing the results of speech recognition and includes acoustic scores, linguistic scores, time points of the beginning end and terminal end of speech portions associated with respective input words;

selection means for selecting one or more candidate first words from the plurality of input words to be processed by speech recognition processing, based on the concatenation information, and a word score that represents an evaluation of acoustic scores and language scores calculated using said characteristic values, and for selecting one or more candidate second words from the plurality of input words not based on the acoustic score, the candidate second words having unstable acoustic characteristic values with a number of phonemes and syllables less than a

preset value, wherein the selection means selects said candidate second words likely to be concatenated linguistically to a directly previously recognized word and to subject the concatenated candidate first and candidate second words to the matching processing;

score calculation means for calculating said score of said concatenated candidate first and candidate second words selected by said selection means referencing concatenation information of said first and second words; and

finalizing means for finalizing a word string, as the recognition result of said speech, based on said score,

wherein the selected one or more candidate first words based on the word score have a number of phonemes and syllables above the number of phonemes and syllables of the candidate second words and more stable acoustic characteristic values than the selected one or more candidate second words,

wherein the word concatenation information is sequentially updated based on the score.

2. – 4. (Canceled)

5. (Previously Presented) The speech recognition apparatus according to claim 1 further comprising:

storage means for memorizing the results of speech recognition;

wherein said selection means selects, as said candidate second words, the input words included in the results of speech recognition memorized in said storage means, with a stored state in said storage means as a measure not based on the non-acoustic score.

6. (Original) The speech recognition apparatus according to claim 5 further comprising:

inputting means for providing an input for correcting the results of speech recognition;
wherein said storage means stores the results of the speech recognition corrected by the input from said inputting means.

7. (Previously Presented) The speech recognition apparatus according to claim 1 wherein said selection means calculates said score using characteristic values of the speech to select said candidate first word based on said score.

8. (Previously Presented) A speech recognition method in which a score reflecting an acoustic likelihood of the results of speech recognition of an input speech is calculated and in which the speech is recognized based on the score, comprising:

an extraction step of extracting characteristic values of said input speech, said input speech comprising a plurality of input words;

a first storing step of storing characteristic values and an extraction time point associated with each characteristic value;

a second storing step of storing word concatenation information that is the relation of input words of a word sequence representing the results of speech recognition and includes acoustic scores, linguistic scores, time points of the beginning end and terminal end of speech portions associated with respective input words;

a selection step of selecting one or more candidate first words from the plurality of input words to be processed by speech recognition processing, based on the concatenation information,

and a word score that represents an evaluation of acoustic scores and language scores calculated using said characteristic values, and for selecting one or more candidate second words from the plurality of input words not based on the acoustic score, the candidate second words having unstable acoustic characteristic values with a number of phonemes and syllables less than a preset value, wherein in the selection step are selected said candidate second words likely to be concatenated linguistically to a directly previously recognized word and to subject the concatenated candidate first and candidate second words to the matching processing;

a score calculation step of calculating said score of said candidate first and candidate second words selected by said selection step referencing concatenation information of said first and second words; and

a finalizing step of finalizing a word string, as the recognition result of said speech, based on said score,

wherein the selected one or more candidate first words based on the word score have a number of phonemes and syllables above the number of phonemes and syllables of the candidate second words and more stable acoustic characteristic values than the selected one or more candidate second words,

wherein the word concatenation information is sequentially updated based on the score.

9. (Currently Amended) A non-transitory computer-readable medium having recorded thereon a program for causing a computer to perform speech recognition processing in which a score reflecting an acoustic likelihood of the results of speech recognition of an input speech is calculated and in which the speech is recognized based on the score, comprising:

an extraction step of extracting characteristic values of said input speech, said input speech comprising a plurality of input words;

a first storing step of storing characteristic values and an extraction time point associated with each characteristic value;

a second storing step of storing word concatenation information that is the relation of input words of a word sequence representing the results of speech recognition and includes acoustic scores, linguistic scores, time points of the beginning end and terminal end of speech portions associated with respective input words;

a selection step of selecting one or more candidate first words from the plurality of input words to be processed by speech recognition processing, based on the concatenation information, and a word score that represents an evaluation of acoustic scores and language scores calculated using said characteristic values, and for selecting one or more candidate second words from the plurality of input words not based on the acoustic score, the candidate second words having unstable acoustic characteristic values with a number of phonemes and syllables less than a preset value, wherein in the selection step are selected said candidate second words likely to be concatenated linguistically to a directly previously recognized word and to subject the concatenated candidate first and candidate second words to the matching processing;

a score calculation step of calculating said score of said candidate first and candidate second words selected by said selection step referencing concatenation information of said first and second words; and

a finalizing step of finalizing a word string, as the recognition result of said speech, based on said score,

wherein the selected one or more candidate first words based on the word score have a number of phonemes and syllables above the number of phonemes and syllables of the candidate second words and more stable acoustic characteristic values than the selected one or more candidate second words,

wherein the word concatenation information is sequentially updated based on the score.